DOCUMENT RESUME

ED 109 918 FL 007 017

AUTHOR Johnson, Bruce C.
TITLE Issues In Sociolinguistic Typology.
PUB DATE 75
NOTE 8p.; Paper presented at the Southeastern Conference on Linguistics (13th, Nashville, Tennessee, March 19-21, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Applied Linguistics; Bilingualism; Data Analysis; Data Collection; Descriptive Linguistics; Diglossia; *Language Classification; Language Planning; Language Research; *Language Typology; *Language Usage; *Language Variation; Multilingualism; Profile Evaluation; *Sociolinguistics

ABSTRACT

Development of a typology for linguistic situations has been a continuing concern of sociolinguistic research. Thus far, however, typological studies have served primarily to demonstrate the difficulties involved in such an undertaking rather than producing results of obvious utility. With few exceptions, notably Greenberg 1956 and Pool 1972, sociolinguistic typology has taken a categorical view of data which is largely scalar. The arbitrary segmentation of scalar attributes leads to undesirable data-loss, but in the absence of a principled measure of significance there is no elegant alternative. Such a measure is to be sought in the study of the ways that sociolinguistic data are applied to practical concerns. In this paper, the general characteristics of sociolinguistic typologies and the particular features of existing models are reviewed and the following major issues are discussed: (1) treatment of scalar attributes and data-loss, and (2) development of objective measures of significance. (Author/AM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED109918

ISSUES IN SOCIOLINGUISTIC TYPOLOGY

This paper is concerned with sociolinguistic typology in the most inclusive sense -- the typology of language situations.[1] Clearly, there are numerous other typologies which may properly be termed sociolinguistic, typologies of language varieties, language functions, speech events, evaluative reactions to language, and so forth, but the typology of language situations holds a pre-eminence over these others in that they are concerned with particular aspects of language situations. Furthermore, while these lower-level typologies are logically prior, ultimately it is the pattern of their inter-relationships that is most significant.

The development of a typology of language situations has been a continuing concern of sociolinguistic research, for good reasons, but thusfar typological studies have served primarily to demonstrate the difficulties involved in such an undertaking, with no clear instances of practical results.

In this paper, I will present briefly the motivations for typological studies in sociolinguistics, review some existing models, and discuss some directions for future research with particular attention to two issues. These issues are data-loss and the need for an objective measure of significance.

FL007017

---

[1] A typology is a set of types which comprise the range of possible combinations of a set of attributes and thereby provides a basis for the classification of the individuals in a universe of discourse. Thus, in a typology of language situations, the universe of discourse comprises language situations, the attributes are characteristics which language situations may have, and individual language situations are of the same type if they display the same combination of attributes.

A language situation is the total configuration of language use in a community, "how many and what kinds of languages are spoken by how many people, under what circumstances, and what the attitudes and beliefs about languages held by the members of the community are" (Ferguson 1966:309).

A community is any natural group of people, in Gumperz' sense of a "human aggregate characterized by regular and frequent interaction over a significant span of time and set off from other such aggregates by differences in the frequency of interaction" (1964:137).

With few exceptions, typological models in sociolinguistics have taken a categorical view of data which is often scalar. To take a familiar example, languages are often stated to be, or not to be, standardized, when in fact all languages are more or less standardized. The arbitrary segmentation of scalar attributes such as standardization results in data-loss which cannot be justified at the current state of the art. (Cf. Pool 1969.)

However, in the absence of a principled measure of significance, there is no elegant alternative to arbitrary segmentation . Such a measure, I suggest, is to be sought in the study of the practical application of sociolinguistic data.

Motivations for sociolinguistic typology

The primary motivation for sociolinguistic typology has been a practical concern with the language-related problems of the developing areas of the world. This practical motivation is based on the acceptance of one or more of the following positions:

a. Description. An adequate description of the significant sociolinguistic characteristics of a community is a useful basis for the development and implementation of language policy;

b. Comparison. The utility of such a description would be increased if it were accompanied by a set of comparable data drawn from a wide range of more and less similar situations;

c. Precision. Most useful of all would be a general and predictive set of rules covering the inter-relationship of linguistic and social processes.

In addition to the practical motivation, there are a number of ways in which typological studies can benefit the field of sociolinguistics. Although these are seldom recognized, they seem equally compelling.

d. Precision. It is not coincidental that the attributes used in typologies are often familiar concepts in the literature. A major part of what is referred to as "developing a typology" is in fact the precision of attributes of an existing informal typology in such a way that they may be unambiguously applied throughout the universe of discourse.

e. Comparability. A rigourous and general typological system can provide a heuristic framework which will promote the collection of comparable data in descriptive studies.

f. Naturalness. An interesting typology should make use of natural classes. That is, types should be established in such a way that individuals of the same type share a number of significant similarities. In this way, the typology will have a certain predictive power, which may in turn serve as a test of validity.

The development of a comprehensive and natural sociolinguistic typology is intimately involved with the development of a general sociolinguistic theory, and is thus of both practical and theoretical interest.

Some models for sociolinguistic typology

At this point I would like to rehearse briefly some existing models for sociolinguistic typology. In so doing, two attributes will be utilized: implicitness and number of types. An implicit typology is one in which the criteria are not clearly defined and appeal is made to common knowledge. Number of types may vary from two to an indefinite number. Larger numbers of types may always be reduced to fewer, but not vice versa.

implicit/few. Ferguson's classic diglossia article (1959) was a contribution to an existing, largely implicit typology of language situations which recognized two main types, monolingual and multilingual. Ferguson established diglossia as a third distinct type. More recently, the creole continuum has, I believe, found a place in this system as a fourth type (DeCamp 1971).

explicit/four. Fishman (1967) has outlined a general classification of language situations based on two binary attributes, the existence of functionally differentiated linguistic repertoires at the individual level, which he calls "bilingualism", and at the community level, which he calls "diglossia". This generates four types of language situations, those characterized by, in his terms, bilingualism with diglossia, bilingualism without diglossia, diglossia without bilingualism, and neither bilingualism nor diglossia.

explicit/indefinite. Greenberg's indexes of linguistic diversity (1956) illustrate the use of scalar as opposed to binary or n-ary attributes. The most useful of the indexes is H, the index of communication, which is "the probability that if two members of the population are chosen at random, they will have at least one language in common" (1956:112). The index ranges from 0 to 1, allowing an indefinite number of types which can of course be reduced in various ways.

explicit/many. Probably the best-known model for sociolinguistic typology is the profile, developed by Stewart (1962, 1968) and Ferguson (1966). The profile is intended as a summary description of the significant aspects of language situations. In practice, this involves primarily three sets of attributes: significance of languages, types of languages, and functions of languages. A language situation is typed by its particular combination of these attributes.

Directions for future research

If the validity of at least some of the motivations for sociolinguistic typology is accepted, there remains the problem of determining the most promising model. I feel that the profile, in the sense of an economical descriptive framework for the significant characteristics of language situations, has the greatest potential practical and theoretical interest of the existing models.

The reasons for this stem from the fact that the profile is more inclusive than the other models, making fewer unwarranted assumptions about the significance of attributes. More exclusive models may easily be derived from more inclusive ones, if desired, but not vice versa. Furthermore, the compilation and presentation of basic sociolinguistic information is of practical interest in itself.

To be sure, there is room for improvement in the profile model. Some current shortcomings of the model involve data-collection, definitions, scope, and the treatment of scalar attributes.

A major problem with profiles is that the relevant data is usually unreliable if available, and often not available. However, if we accept the importance of sociolinguistic data, then collection difficulties must be faced. In fact, one motivation for the development of the profile model was that it would stimulate the collection of relevant sociolinguistic information.

It is essential that definitions of attributes be explicit and empirical. Too much has been left too often to the discretion of the individual investigator, thus promoting uncertainty and inconsistency.

Scope refers to the delimitation of the universe of discourse. It seems clear that the universe of discourse for sociolinguistic typology should comprise language situations of communities, as defined above, rather than of nations or politics, as in the past. In the first place, nations consist of communities which may differ widely in their sociolinguistic patterns. Community studies will necessarily include features which are national, but the reverse is not true. Secondly, the most extensive and reliable sociolinguistic data available are to be found in studies of subnational units such as Jersey City, Khalapur, and the Lower East Side.

Data-loss, the neglect or waste of significant information already at hand, has already been mentioned in several contexts. In particular, the profile model was preferred because it is relatively conservative of data. Similarly, communities were preferred over nations in the discussion of the scope of typologies to further reduce data-loss. Another source of data-loss is the arbitrary segmentation of scalar attributes. If a typological model requires that, before being used, continuous data be compressed into a di- or trichotomy, potentially significant data is lost. This is of frequent occurrence with the profile model.

For example, the proportion of native speakers of a language variety in a community may range over an indefinite number of values between 0 and 1. Ferguson's profile model requires that this continuous variable be trichotomized. If a language is spoken natively by more than twenty-five percent of the population of the community it is a major language, if by between five and twenty-five percent it is a minor language, and if by less than five percent it is excluded from the profile.[1] The precision of this definition is to be applauded. Given the relevant data, it can be easily applied and the number of major and minor languages in the community unambigouusly and consistently determined. However, what is the justification for this trichotomy? Why not segment at six and twenty-six, or at ten and thirty? Other cutting points have in fact been suggested. Stewart (1968), for instance, would segment the variable into six sections, but still without convincing justification. Clearly, a profile which includes the actual proportions would be more informative.

For another example, consider relatedness of language varieties. Here again is a continuous variable with values ranging from 0, unrelated, to 1, identical. This particular attribute of languages was not originally included in the profile model by Ferguson (1966) or Stewart, but has been included by Ferguson (1970) and others, e.g. Roberts 1962. In Ferguson (1970), five degrees of relationship are distinguished. Language varieties may be in different families, in the same family, in the same branch of a family, in a diglossic relationship within the same language, or in the same language and therefore not distinguished. Again, the cutting points are reasonably explicitly defined, but not justified, and potentially significant information is omitted.

---

[1] This is a simplified presentation of this aspect of Ferguson's model. Proportion of native speakers is only one of several criteria for rating the significance of languages. A language may have no native speakers in the community and still be a major language, by another criterion. The argument made here is, however, equally applicable to these other criteria.

The justifications for omission of data are in terms of descriptive economy. It would be impossible to present a totally inclusive description of the language situation in a community, or even to collect all of the necessary data to base one on. Comparison, especially, is facilitated by conciseness. Furthermore, not all data is equally valuable and the scientific investigator must be discriminating in this regard. Finally, typologies by definition treat distinct individuals as the same. We hardly have a typology if each individual in the universe of discourse is of a different type.

The fact is, however, that descriptive economy should be achieved by the principled omission of non-significant or redundant information, rather than by arbitrary exclusion. A profile which includes too much is to be preferred over one which excludes too much. That is, minimization of data-loss must take precedence over descriptive economy. Of course, the optimal profile will include all and only the significant characteristics of a language situation, and the optimal typological model will specify which characteristics these are. That is, it will provide an objective, meaningful measure of significance, a measure which is not presently at hand.

I suggest that such a measure should be sought in the study of the ways in which sociolinguistic data in general, and comparative sociolinguistic data in particular, are applied to practical concerns, a study that should be undertaken in any event.

How are the results of descriptive sociolinguistic studies relevant to practical concerns? The assumption in sociolinguistic typology is that this involves the comparison of the results of particular language policies in particular language situations. More specifically, it is that if a particular policy has had certain results in one situation, the same policy should have the same results in another situation of the same type.

This provides, I believe, the necessary measure of significance for typological studies. Thus, if the same policy has different results in two situations of the same type, then there is a significant difference between the situations and they should not have been classed together. The classification error may have been due to incorrect descriptive data or to a flaw in the profile framework. In any event, it would at least be clear that there was an error and its source investigated.

Adoption of such a practical definition of significance would greatly improve the practical utility of profiles at the same time as providing a principled basis for descriptive economy and a more sophisticated approach to basic sociolinguistic interests.

References

DeCamp, David. 1971. Toward a generative analysis of a post-creole continuum. Pidginization and creolization of languages, ed. by Dell Hymes, pp. 349-370. Cambridge: Cambridge University Press.

Ferguson, Charles A. 1959. Diglossia. Word 15.325-340.

_____. 1966. National sociolinguistic profile formulas. Sociolinguistics, ed. by William Bright, pp. 309-324. The Hague: Mouton.

_____. 1970. The role of Arabic in Ethiopia: A sociolinguistic perspective. Georgetown University Monograph Series on Languages and Linguistics 23.355-368.

Fishman, Joshua A. 1967. Bilingualism with and without diglossia; diglossia with and without bilingualism. Journal of Social Issues 23.2.29-38.

Greenberg, Joseph H. 1956. The measurement of linguistic diversity. Language 32.109-115.

Gumperz, John J. 1964. Linguistic and social interaction in two communities. American Anthropologist 66,6,2.137-153.

Pool, Jonathan. 1969. National development and linguistic diversity. La Monda Lingvo-Problemo 1.129-192.

Rice, Frank A. ed. 1962. Study of the role of second languages in Asia, Africa, and Latin America. Washington: Center for Applied Linguistics.

Roberts, Janet. 1962. Sociocultural change and communication problems. In Rice 1962:105-123.

Stewart, William A. 1962. An outline of linguistic typology for describing multilingualism. In Rice 1962:15-25.

_____. 1968. A sociolinguistic typology for describing national multilingualism. Readings in the sociology of language, ed. by Joshua A. Fishman, pp. 530-545. The Hague: Mouton.